MONTE CARLO METHODS FOR LATTICE FIELDS

by

Richard S. Crosby

SUBMITTED TO THE DEPARTMENT OF
ELECTRICAL ENGINEERING AND COMPUTER
SCIENCE IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF SCIENCE IN ELECTRICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1989

Signature of Author _____
Department of Electrical Engineering and Computer Science
May 22, 1989

Certified by _____
Professor Sanjoy K. Mitter
Thesis Supervisor

Accepted by _____
Leonard A. Gould
Chairman, Department Committee on Undergraduate Theses

MONTE CARLO METHODS FOR LATTICE FIELDS

by

Richard S. Crosby

Submitted to the Department of Electrical Engineering and Computer Science on May 22, 1989 in partial fulfillment of the requirements for the degree of Bachelor of Science in Electrical Engineering.

Abstract

A discussion of the Monte Carlo methods for the Ising model highlights the benefits of the Swendsen-Wang algorithm. Next, an a priori model that has been developed for the analytical reconstruction of crystal structures from x-ray crystallography data is discussed, and the feasibility of developing an algorithm for this model similar to the Swendsen-Wang algorithm is considered.

Thesis Supervisor: Professor Sanjoy K. Mitter
Title: Professor of Electrical Engineering and Computer Science

To Matthew

Table of Contents

List of Figures

Chapter 1

Introduction

Monte Carlo methods are essential tools for solving probabilistic problems involving lattice fields. This work explores and discusses the use of Monte Carlo methods in the context of lattice fields and intends to propose new ideas for solving problems that are currently being examined. First, we discuss the Ising model, explaining what can be solved with it and what cannot be solved. Next, we consider three algorithms - the Metropolis, Gibbs sampler, and Swendsen-Wang algorithms - and after a qualitative discussion of the strengths and weaknesses of each algorithm, we compare experimental results obtained for each of the algorithms and draw conclusions about the rapidity of convergence for each of the algorithms. Then, we discuss an a priori model for determining crystal structures from x-ray crystallography data. Finally, we discuss the possibility of creating an algorithm for the crystallography model that would be analogous to the Swendsen-Wang algorithm.

Chapter 2

The Ising Model

2.1 Background

As stated in Kindermann and Snell (1980), the Ising model was originally developed by Ernst Ising as his doctoral thesis project under the supervision of Lenz. This model was used in an attempt to explain certain empirically observed facts about ferromagnetic materials. Since the inception of the model, it has also been found to be applicable to modeling other systems that are not necessarily physical, such as gases, binary alloys, cell structures, and even sociological models.

2.2 Description

The Ising model consists of variables positioned at sites on a lattice L. These site variables can have one of two values at any point in time, the values usually consisting of +1 and -1. Let us define Ω to be the sample space of all possible configurations ω that can be formed by the site variables. Next, consider the configuration energy function

$$U(\omega) = -J \sum_{<i,j>} \omega_i \omega_j \qquad (2\text{-}1)$$

where ω_n is the variable for site n in a configuration ω which takes on a value of either +1 or -1, and J is a coupling constant that determines the strength of the interactions and is usually assumed to be 1. The summation is

over all pairs of sites i and j that are separated by one unit distance. Using this configuration energy function, we can define the invariant probability measure

$$\pi = \frac{1}{Z} exp(-U/T) \tag{2-2}$$

where T is a temperature parameter of the system and Z is a normalizing constant. Thus, the probability of any configuration can be determined from equation 2-2.

As stated by Marroquin (1985) and Kindermann and Snell (1980), the Ising model is a Markov random field. A Markov random field is defined as a collection of random variables corresponding to the sites on a lattice which has a probability distribution that satisfies the following relation:

$$P(\omega_j=q|\omega_k,k\neq j) = P(\omega_j=q|\omega_k,k\varepsilon N_j) \tag{2-3}$$

where j is an arbitrary site on the lattice, N_j is the set of all the neighbors of site j, and q is an arbitrary value for the variable at site j. Essentially, sites are only influenced by other sites within its neighborhood while it is not influenced at all by sites outside its neighborhood.

The entropy of any measure π on a sample space Ω, as stated by Kindermann and Snell (1980), is defined as

$$S(\pi) = -\sum_{\omega} \pi(\omega)\, log\pi(\omega) \qquad (2\text{-}4)$$

where the summation is over all configurations ω in the sample space Ω. The invariant measure in equation 2-2 is the measure that maximizes the entropy out of all valid measures given a certain expected energy.

The Ising model is characterized by three temperature regions: below the critical temperature, near the critical temperature, and above the critical temperature. As stated by Marroquin (1985) the critical temperature is defined to be the maximum temperature for which fixed conditions on the boundary of a square lattice are felt at the center of the lattice no matter how large the lattice is.

2.3 Onsager's Solution

Huang (1963) gives a full description of Onsager's solution to the two -dimensional Ising model. This solution is of great importance since it provides a method for determining properties of the Ising model analytically rather than experimentally and is thus an essential tool when using the ising model. Some of the physical properties solved for including the internal energy per spin

$$u = -coth\tfrac{2}{T}[1 + \tfrac{2}{\pi}\kappa' K_1(\kappa)] \qquad (2\text{-}5)$$

where

$$\kappa = 2sinh\frac{2}{T}/cosh^2\frac{2}{T} \tag{2-6}$$

and

$$\kappa' = 2tanh^2\frac{2}{T} - 1 \tag{2-7}$$

and $K_1()$ is the complete elliptical integral of the first kind, and the specific heat

$$c = \frac{2}{\pi}(\frac{1}{T}coth\frac{1}{T})^2\{2K_1(\kappa) - 2E_1(\kappa) - (1-\kappa')[\frac{\pi}{2} + \kappa' K_1(\kappa)]\} \tag{2-8}$$

where $E_1()$ is the complete elliptical integral of the second kind. The critical temperature can be derived from these formulas and the actual formula for the critical temperature is

$$T_c = 2 / sinh^{-1}1 \tag{2-9}$$

2.4 Limitations of the Model

There are limits to what can be solved with the Ising model. Since a probability measure is typically assigned to a sample space which represents unobservable outcomes, only very broad properties of the system can be observed, such as the properties determined by Onsager.

Chapter 3

Monte Carlo Algorithms - The Ising Model

Many Monte Carlo algorithms have been developed for generating sample functions of Markov random fields. However, only three algorithms - Metropolis, Gibbs sampler, and Swendsen-Wang - were chosen for tests involving the Ising model for this project. The Metropolis algorithm was chosen since it represents the first successful algorithm and was also the simplest in its approach. The Gibbs sampler was chosen since its deterministic approach in its selection of sites to be visited presents itself well for a parallel processing approach. Finally, the Swendsen-Wang algorithm was selected since it represents a unique approach to overcoming the problem of critical slowing down.

3.1 The Metropolis Algorithm

3.1.1 Description

The methodology behind the algorithm proposed by Metropolis is as follows: First, choose a site j from the lattice at random. Let the current configuration be ω_0 and the current state of site j be ω_{j0}. Next, randomly pick a new state ω_{j1} for site j from the set of allowable states with a uniform probability distribution. Then, compute the change in energy ΔU that results from changing the state of j from ω_{j0} to ω_{j1}. If ΔU is less than or equal to zero, then change the state of j to ω_{j1}. Otherwise, generate a random number r uniformly distributed between zero and one. If the r is less than or equal to

exp(-ΔU/T), then change the state of j to ω_{j1}; otherwise, leave j unchanged. Finally, repeat the steps for enough iterations until convergence is obtained.

3.1.2 Proof of Reversibility

We wish to show that

$$\pi(\omega_0)P(\omega_0,\omega_1) = \pi(\omega_1)P(\omega_1,\omega_0) \tag{3-1}$$

where π is the invariant measure defined in the preceding chapter, P is the transition matrix for the Markov chain, and ω_0 and ω_1 are two configurations in the sample space Ω. This equation is known as the detailed balance equation. Using the notation from the description given above, we have

$$P(\omega_0,\omega_1) = \frac{1}{N} min[1\,,\, exp(-\Delta U/T)] \tag{3-2}$$

and

$$P(\omega_1,\omega_0) = \frac{1}{N} min[1\,,\, exp(\Delta U/T)] \tag{3-3}$$

where N represents the total number of sites on the lattice L. Hence, since

$$\pi(\omega_1) = \pi(\omega_0)exp(-\Delta U/T) \tag{3-4}$$

we have

$$\pi(\omega_0)\tfrac{1}{N} = \pi(\omega_0)exp(-\Delta U/T)\tfrac{1}{N}exp(\Delta U/T) \tag{3-5}$$

and

$$\pi(\omega_0)\tfrac{1}{N}exp(-\Delta U/T) = \pi(\omega_0)exp(-\Delta U/T)\tfrac{1}{N} \tag{3-6}$$

where equation 3-5 represents the case when ΔU is less than or equal to zero and equation 3-6 represents the case when ΔU is greater than zero. Thus, since both equation 3-5 and equation 3-6 are correct, equation 3-3 is satisfied for all ω_0 and ω_1. Therefore, the Markov chain is reversible, and hence, the algorithm statisfies the invariant measure.

3.2 The Gibbs Sampler Algorithm

3.2.1 Description

The Gibbs sampler algorithm as proposed by Geman and Geman (1984) is as follows: Sites in the lattice L are visited in some deterministic order. At a given site j, choose a new site at random from the conditional distribution

$$P(\omega_{j1} = q \mid \omega_0) = exp(-\Delta U_q/T)/\textstyle\sum_{p\varepsilon Q} exp(-\Delta U_p/T) \tag{3-7}$$

where ΔU_q is the change in configuration energy if site j takes on state q and Q is the set of all possible states. Repeat these steps for all sites in the lattice for enough iterations to obtain satisfactory results.

If we simplify equation 3-7 to consider a two-state system such as the Ising model, and defining ΔU to be the change in configuration energy if site j were to take on the state other than its current state. We then have

$$P(\omega_1=\omega_0 \mid \omega_0) = 1/[1+exp(-\Delta U/T)] \tag{3-8}$$

and

$$P(\omega_1\neq\omega_0 \mid \omega_0) = 1/[1+exp(\Delta U/T)] \tag{3-9}$$

which is known as the heat bath algorithm for a binary system.

3.2.2 Proof of Reversibility

Once again we wish to show that equation 3-3 holds for arbitrary configurations ω_0 and ω_1. We will limit this proof to arbitrary two-state systems such as the Ising model. Defining ΔU to be the difference in configuration energy between ω_0 and ω_1 and using equations 3-4, 3-8, and 3-9, we get

$$\pi(\omega_0)\frac{1}{1+exp(\Delta U/T)} = \pi(\omega_0)exp(-\Delta U/T)\frac{1}{1+exp(-\Delta U/T)} \tag{3-10}$$

which is correct. Thus, the Gibbs sampler algorithm is reversible for at least two-state systems, and this proof can be extended to more general systems.

3.3 The Swendsen-Wang Algorithm

3.3.1 Description

Swendsen and Wang (1987) derived their Swendsen-Wang algorithm as follows: First, consider the configuration energy function for a Potts model

$$U = K\sum_{<i,j>}(1-\delta_{\sigma_i,\sigma_j}) \tag{3-11}$$

where the temperature parameter has been absorbed into K. From this function we define the invariant measure to be

$$\pi = \frac{1}{Z}exp(-U/T) \tag{3-12}$$

where Z is a normalizing constant. Next, remove the energy contribution due to the interaction between sites l and m from the energy function. Then we have

$$U_{l,m} = K\sum_{<i,j>\neq<l,m>}(1-\delta_{\sigma_i,\sigma_j}) \tag{3-13}$$

Using this function we can define the two restricted sums

$$\pi_{l,m}^{same} = \frac{1}{Z}exp(-U_{l,m}/T)\delta_{\sigma_l,\sigma_m} \tag{3-14}$$

and

$$\pi_{l,m}^{diff} = \frac{1}{Z} exp(-U_{l,m}/T)(1-\delta_{\sigma_l,\sigma_m}) \tag{3-15}$$

Using these terms we can rewrite equation 3-12 as

$$\pi = \pi_{l,m}^{same} + exp(-K)\pi_{l,m}^{diff} \tag{3-16}$$

Next, we define a new term representing the invariant measure independent of the interaction of sites *l* and *m*:

$$\pi_{l,m}^{ind} = \pi_{l,m}^{same} + \pi_{l,m}^{diff} \tag{3-17}$$

Finally, rewriting equation 3-16 in terms of equation 3-17, we get

$$\pi = [1-exp(-K)]\pi_{l,m}^{same} + exp(-K)\pi_{l,m}^{ind} \tag{3-18}$$

Since the second term in equation 3-18 is independent of the interaction between sites *l* and *m* and the first term restricts the spin variables at sites *l* and *m* to be the same, the probability, $p = 1 - \exp(-K)$, can be interpreted as a bond between the two sites.

Swendsen and Wang make use of this bond nature in the formulation of their algorithm for the Potts model. The methodology of the algorithm is as follows: First, visit every nearest neighbor interaction between sites *i* and *j* on the lattice. If sites *i* and *j* have the same state, then a bond is generated with probability $p = 1 - \exp(-K)$; otherwise, no bond is formed. This process

continues until all nearest neighbor interactions have been visited. Thus, we are left with a bond configuration, upon which we define clusters which consist of sites on the lattice connected by bonds. Now, for each of the clusters, we randomly assign with a uniform distribution a new state out of the set of possible states and then assign this state to each site on the lattice in the cluster. The process is then repeated until enough iterations have been performed to achieve satisfactory results.

3.3.2 Proof of Reversibility

First, we note that if q is the number of possible Potts states and N_c is the number of clusters, then the invariant measure can be rewritten as

$$\pi = \frac{1}{Z} p^b (1-p)^n q^{N_c} \tag{3-19}$$

where b is the number of bonds and n is the number of interactions with the same Potts state that did not form a bond. The probability of passing through a particular bond/cluster configuration has a factor of p for each bond and a factor of q for each cluster, but differs by a number of factors of $(1-p) = \exp(-K)$ for each interaction with the same state that did not form a bond. The probability of going from a bond/cluster configuration to a Potts configuration is uniform, so the bond/cluster configuration uniquely describes the resulting Potts configuration. Thus, the two Potts configurations differ in configuration energy by a total of Δn terms of K, where Δn is the difference of the number of interactions with the same state that did not form a bond for two arbitrary Potts configurations, ω_0 and ω_1, which pass through the same bond/cluster configuration. So, we have

$$\pi(\omega_1) = \pi(\omega_0) exp(-\Delta nK) \tag{3-20}$$

Finally, applying equations 3-19 and 3-20 to equation 3-1, we get

$$\pi(\omega_0) \frac{1}{Z} p^b (1-p)^{n+\Delta n} q^{N_c} = \pi(\omega_0) exp(-\Delta nK) \frac{1}{Z} p^b (1-p)^n q^{N_c} \tag{3-21}$$

which is a correct equation. Therefore, the Swendsen-Wang algorithm produces a reversible chain.

3.4 Qualitative Comparison of the Algorithms

All three algorithms discussed above have about the same execution efficiency in terms of average time required to update a site. The Metropolis algorithm is the most straightforward of the three algorithms to implement. However, the Metropolis algorithm has no unique feature that allows it to achieve greater efficiency, unlike the other two algorithms which do. The Gibbs sampler algorithm is the next easier to implement, but can require many calculations for each iteration, including calculations of exponentials. Nevertheless, the Gibbs sampler is made worthwhile by the fact that it can be implemented in a parallel processing fashion since it is deterministic in its choice of next site to be updated and the Markovian nearest neighbor property isolates the sites to a certain extent. The only necessary condition for parallel implementation of the Gibbs sampler is that two neighbors never try to update at the same time. The Swendsen-Wang algorithm is difficult to implement efficiently. However, it is possible to implement the algorithm so that the average access time per site on the lattice is about the same as that

of the Metropolis algorithm. The main benefit of the Swendsen-Wang algorithm lies in its inherent ability to achieve convergence with significantly fewer iterations than the other algorithms.

Chapter 4

Experimental Results - The Ising Model

4.1 Introduction

The three algorithms described in the preceding chapter - Metropolis, Gibbs sampler, and Swendsen-Wang - were implemented on an Ardent Titan computer, and experimental data were generated for a number of different test cases. The main goals for these experiments were to get a feel for the nature in which the algorithms approached their solutions and to obtain comparative information between the algorithms for different parameters and conditions. The comparative information included relative efficiency, effort required to achieve acceptable accuracy with respect to theoretical results, and overall performance.

4.2 Equilibrium Energy

The first major block of experiments was concerned with determining lattice equilibrium energy per spin for each of the algorithms in each of the three temperature regions. Conditions such as different lattice sizes and numbers of iterations were also considered in these trials. However, all trials were conducted with free boundary conditions (no interaction at lattice edges), and the initial state of the lattice for each of the trials was an entirely random configuration.

4.2.1 Low Temperature

The first group of trials for this block of experiments had the following conditions. The trials were run for all three algorithms operating in the low temperature region (at about half the critical temperature), and results for three lattice sizes - 32x32, 64x64, and 128x128 - were considered. Figure 4-1 on page 23 shows the results of a sample of one hundred trials for each set of conditions with an execution length of one hundred iterations per lattice site. Figure 4-2 on page 24 show these results for samples of ten trials with an execution length of one thousand iterations per lattice site.

As can be seen in both figure 4-1 and figure 4-2, the Swendsen-Wang algorithm approaches the theoretical result more rapidly than the other algorithms. Since the Metropolis and Gibbs sampler algorithms approach the equilibrium energy asymptotically with a slope that gets flatter with respect to time as the temperature decreases, these algorithms require a large increase in the number of iterations to get a small improvement in the difference between the theoretical and experimental values for the equilibrium energy. The Swendsen-Wang algorithm overcomes this sluggishness through its ability to alter large portions (clusters) of the lattice in a single iteration as discussed in the preceding chapter.

4.2.2 Critical Temperature

The next group of trials concerned with equilibrium energy consisted of trials with the temperature fixed near the critical temperature. These trials were also run for all three algorithms for each of the three lattice sizes. Figure 4-3 on page 26 shows the results for one hundred samples of one hundred

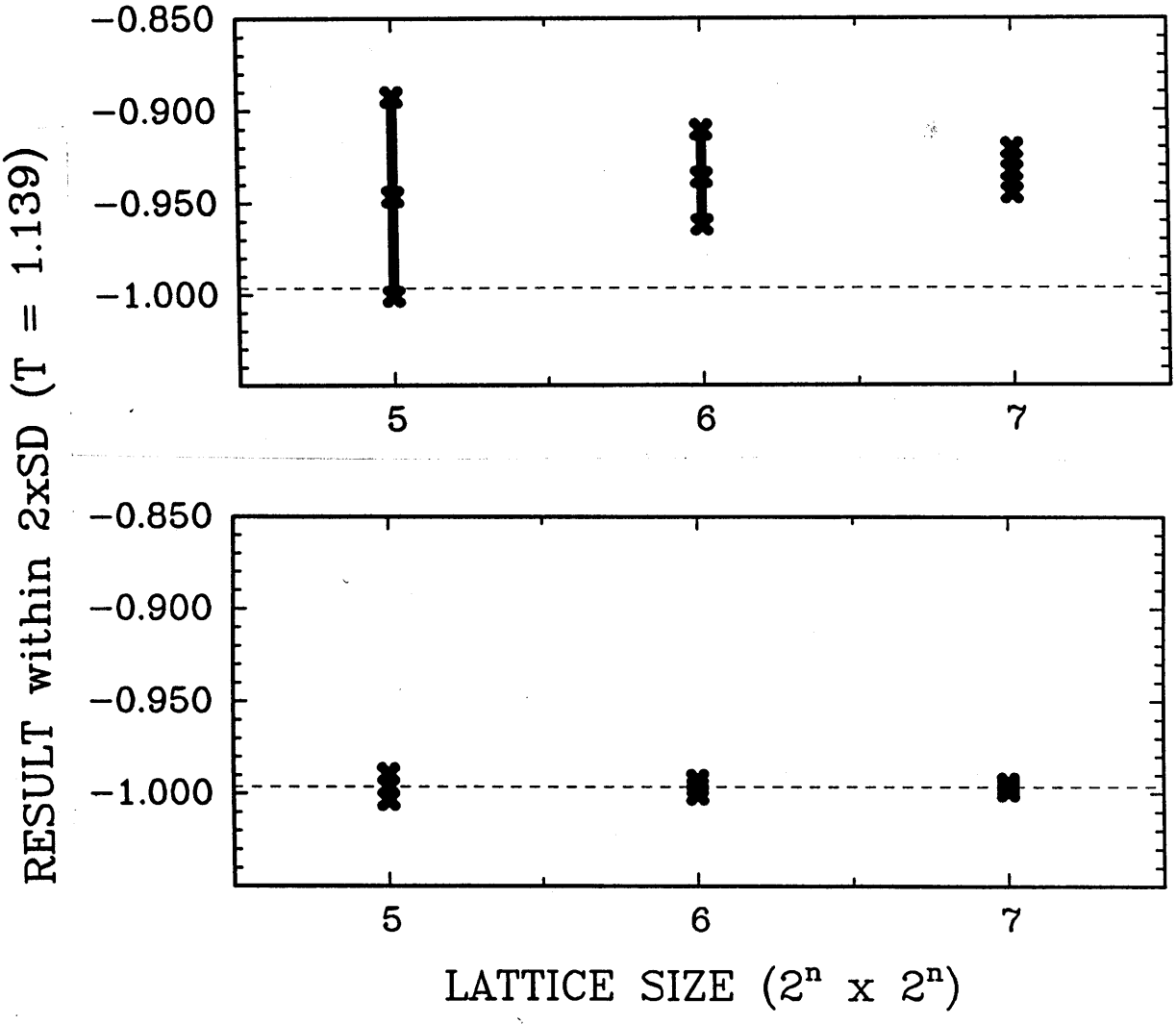

Figure 4-1: Equilibrium Energy - Low Temperature

100 samples, 100 iterations per lattice site. The dotted line corresponds to the value calculated by analytical methods. The bars correspond to the results obtained for the algorithm within two standard deviations of the mean. Algorithms from top to bottom: Metropolis, Swendsen-Wang. Graphical data for the Gibbs sampler were not available.

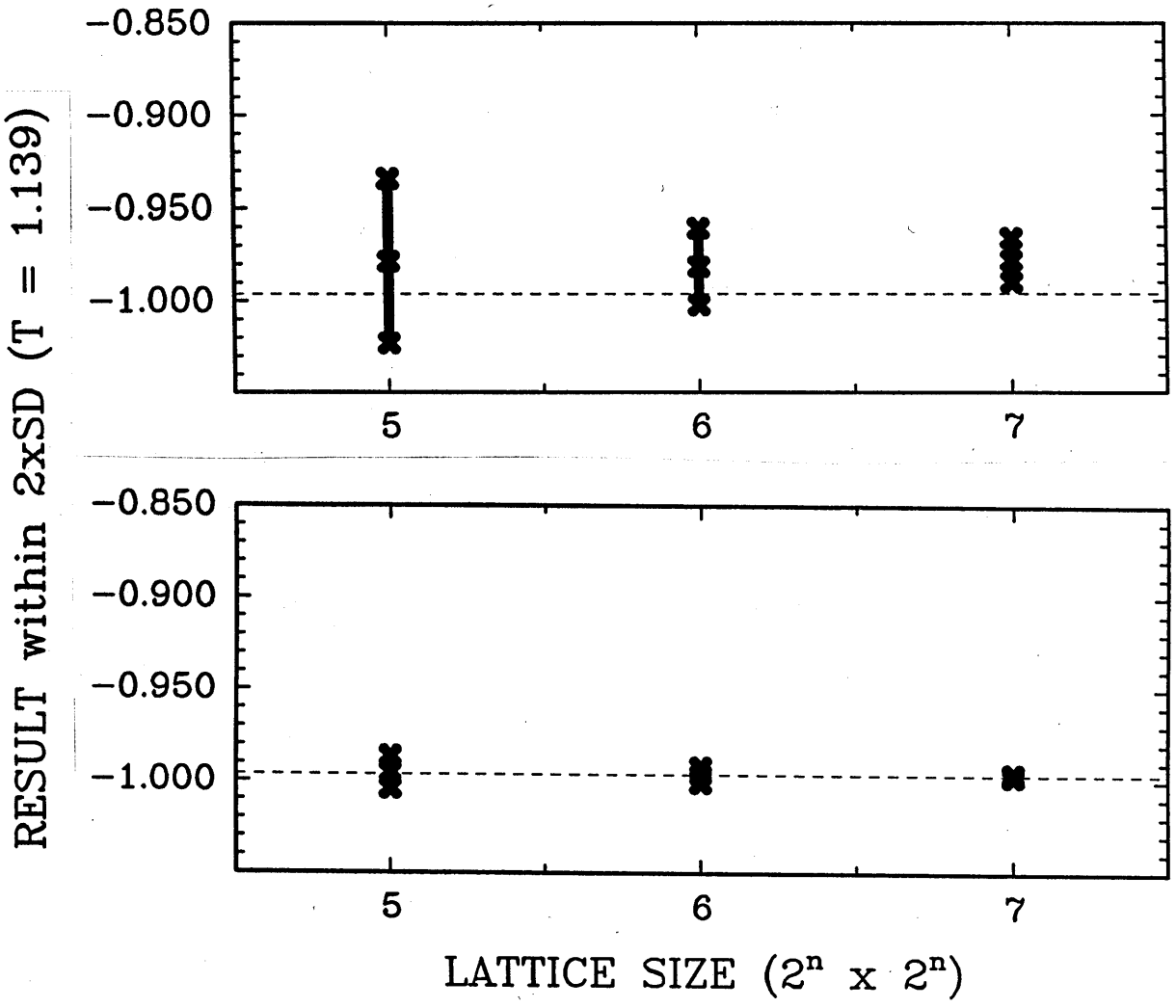

Figure 4-2: Equilibrium Energy - Low Temperature

10 samples, 1000 iterations per lattice site. The dotted line corresponds to the value calculated by analytical methods. The bars correspond to the results obtained for the algorithm within two standard deviations of the mean. Algorithms from top to bottom: Metropolis, Swendsen-Wang. Graphical data for the Gibbs sampler were not available.

iterations per site, and figure 4-4 on page 27 shows the results for ten samples of one thousand iterations per site.

Once again, the contrast between the rapid approach of the Swendsen-Wang algorithm to the equilibrium energy as compared to the slow progress of the Metropolis and Gibbs sampler algorithms is immediately evident in the two figures. Another interesting observation is the difference in the rates of convergence of the three lattice sizes. In general the smaller the lattice spacing (ie. the larger the lattice), the quicker the energy converges to the equilibrium value. This effect can be attributed to the finite size of the lattices. The invariant measure is actually defined for an infinite lattice, so the larger the lattice spacing, the more significant the "edge effects" are upon the results of the trials. It is interesting to note that these effects are most pronounced when the temperature is close to the critical temperature.

4.2.3 High Temperature

The last group of trials concerned with equilibrium energy per site consisted of trials with a temperature much higher than the critical temperature. Trials were run for each of the three algorithms with each of the three lattice sizes. Figure 4-5 on page 28 shows the results of the trials for one hundred samples of one hundred iterations per site while figure 4-6 on page 29 shows the results for ten samples of one thousand iterations per site.

As can be observed from the two figures, all of the algorithms converge to the equilibrium energy quickly with little or no difference resulting from lattice size or choice of algorithm. In terms of performance, the Swendsen

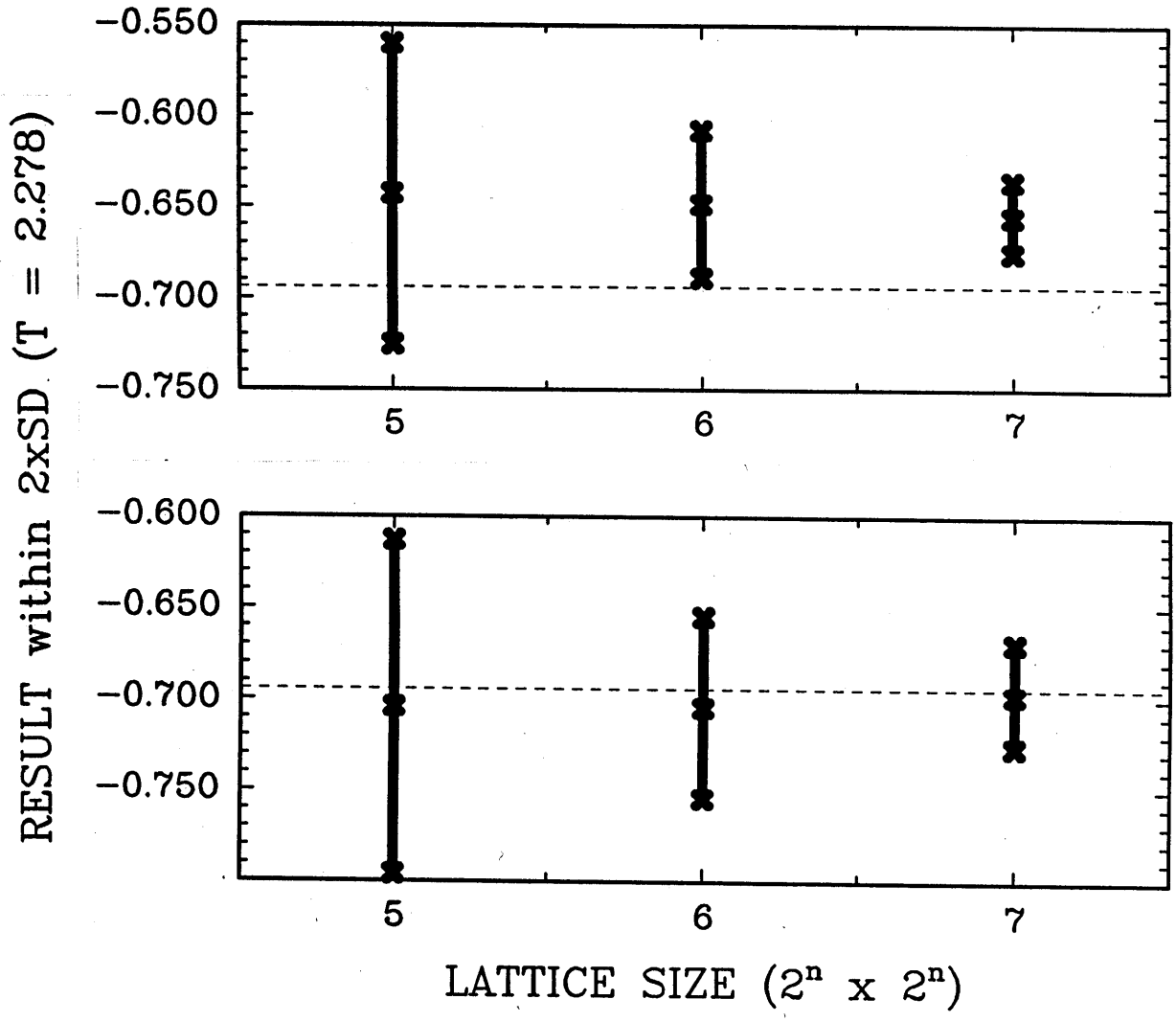

Figure 4-3: Equilibrium Energy - Critical Temperature

100 samples, 100 iterations per lattice site. The dotted line corresponds to the value calculated by analytical methods. The bars correspond to the results obtained for the algorithm within two standard deviations of the mean. Algorithms from top to bottom: Metropolis, Swendsen-Wang. Graphical data for the Gibbs sampler were not available.

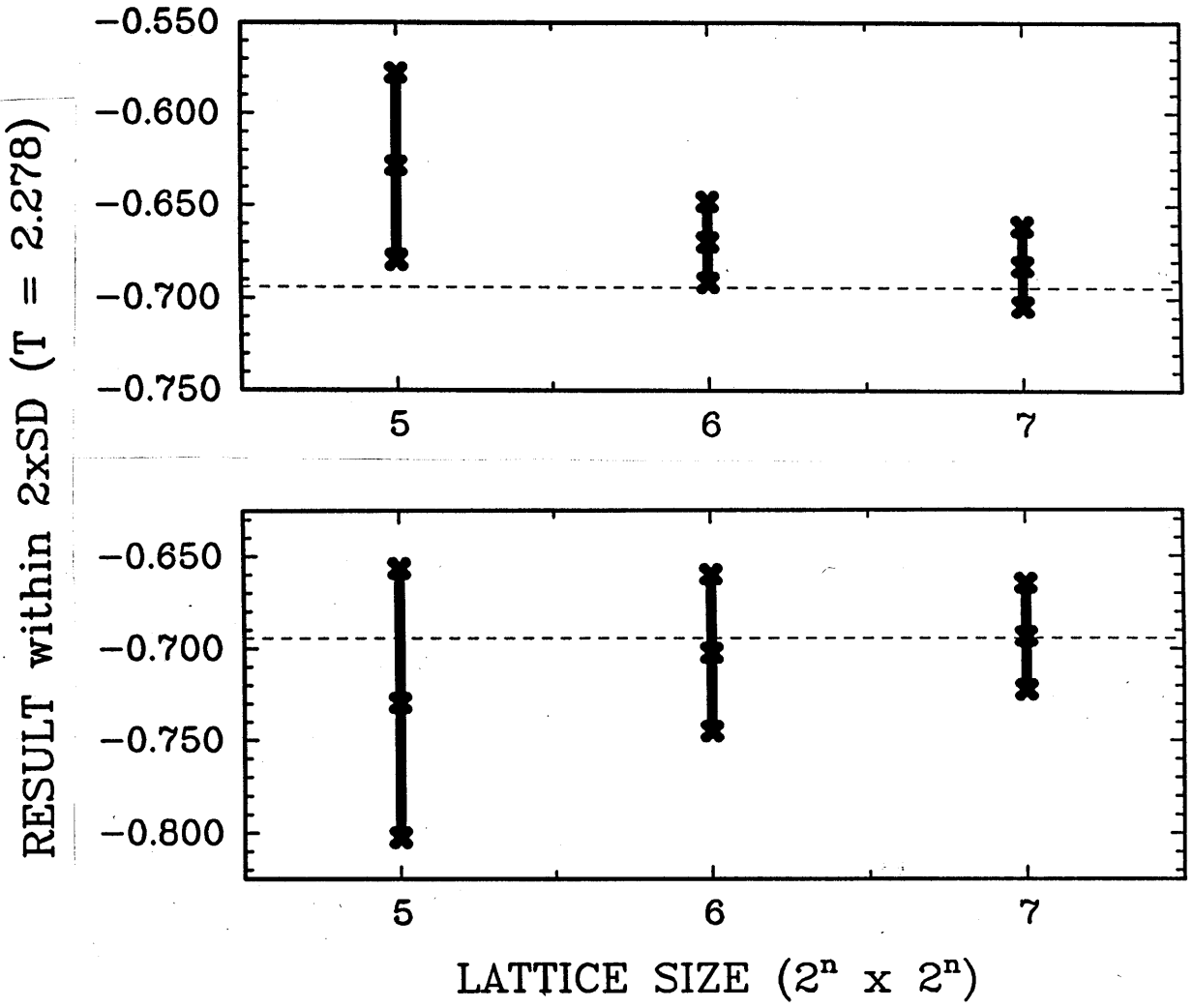

Figure 4-4: Equilibrium Energy - Critical Temperature

10 samples, 1000 iterations per lattice site. The dotted line corresponds to the value calculated by analytical methods. The bars correspond to the results obtained for the algorithm within two standard deviations of the mean. Algorithms from top to bottom: Metropolis, Swendsen-Wang. Graphical data for the Gibbs sampler were not available.

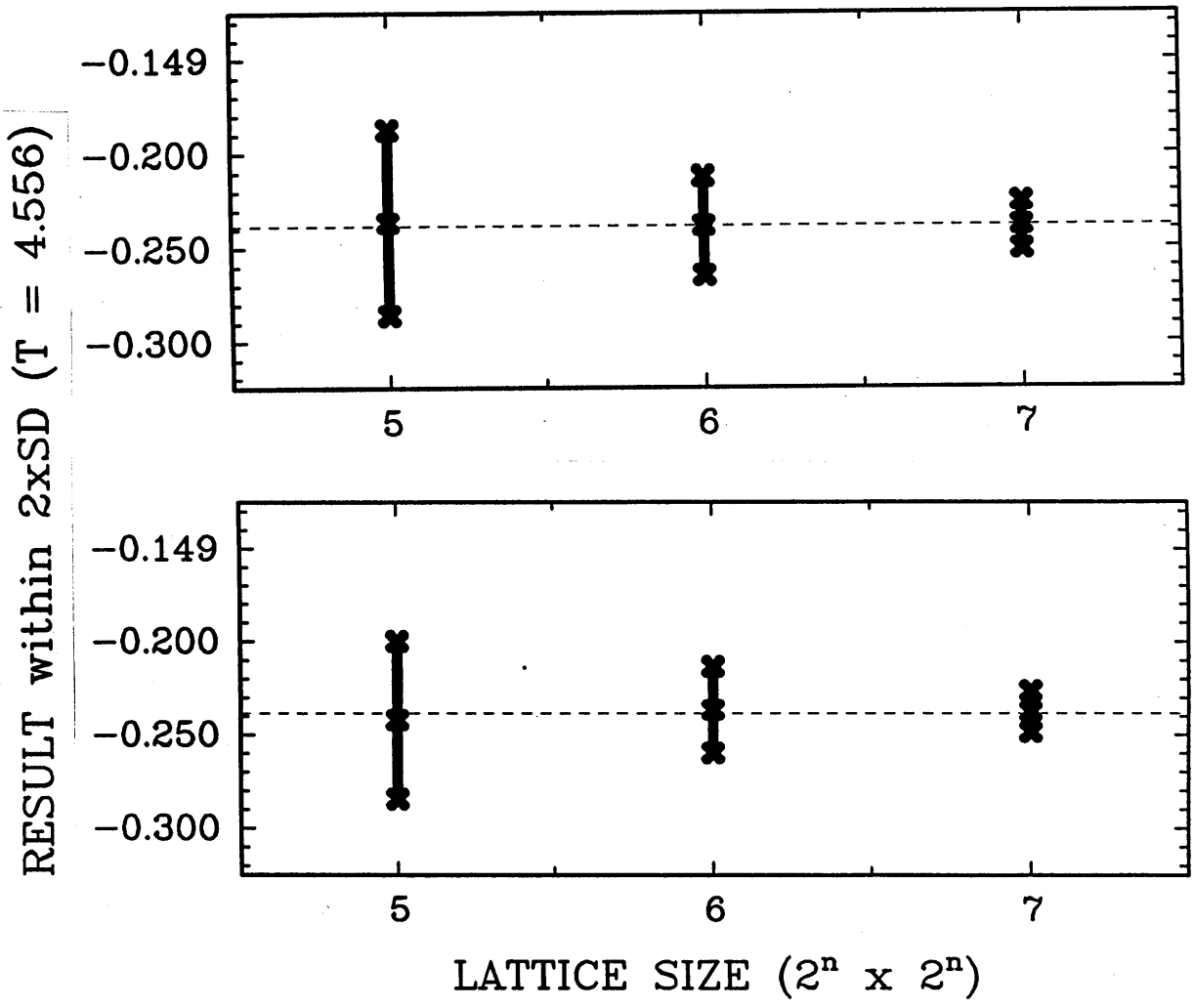

Figure 4-5: Equilibrium Energy - High Temperature

100 samples, 100 iterations per lattice site. The dotted line corresponds to the value calculated by analytical methods. The bars correspond to the results obtained for the algorithm within two standard deviations of the mean. Algorithms from top to bottom: Metropolis, Swendsen-Wang. Graphical data for the Gibbs sampler were not available.

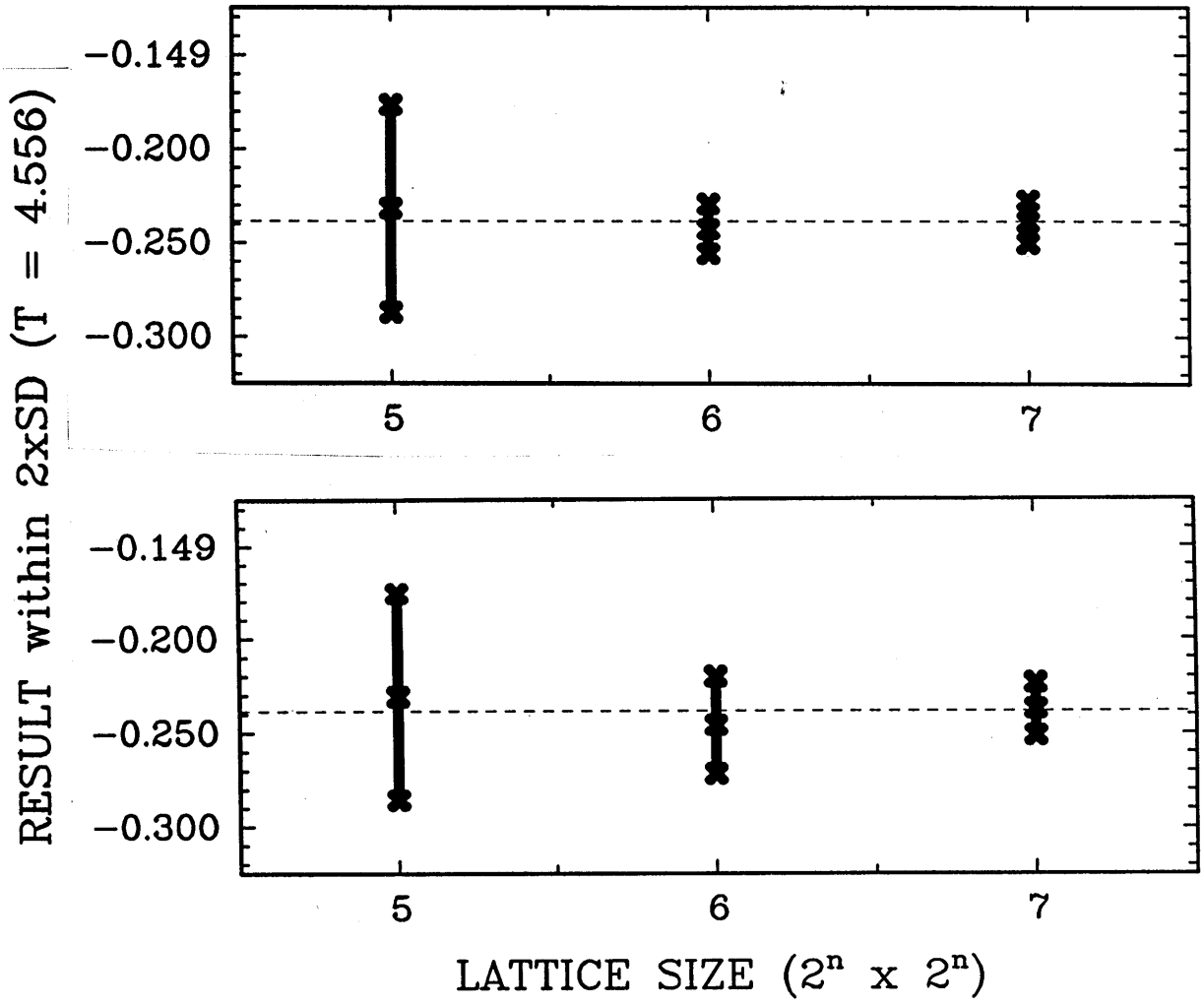

Figure 4-6: Equilibrium Energy - High Temperature

10 samples, 1000 iterations per lattice site. The dotted line corresponds to the value calculated by analytical methods. The bars correspond to the results obtained for the algorithm within two standard deviations of the mean. Algorithms from top to bottom: Metropolis, Swendsen-Wang. Graphical data for the Gibbs sampler were not available.

-Wang algorithm becomes almost identical to the Metropolis and Gibbs sampler algorithms in this temperature range because most clusters consist of only single sites and thus the ability to manipulate large portions of the lattice in a single iteration is lost.

4.3 Temporal Correlation of Equilibrium Energy over Time

The second major block of experiments was concerned with determining the temporal correlation of the equilibrium energy over time for each of the algorithms in each of the three temperature regions. Each of the trials was conducted upon a 64x64 lattice with periodic boundary conditions. The initial state of the lattice for each of the trials was an entirely random configuration. The trials were run for both one hundred iterations per site and one thousand iterations per site.

4.3.1 Low Temperature

The first set of trials for this block of experiments was conducted in the low temperature region. Figure 4-7 on page 31 shows the results for each of the three algorithms for a run of one hundred iterations per site. Figure 4-8 on page 32 shows the same information except for a run of one thousand iterations per site.

The most notable feature of these figures is the marked difference between the Swendsen-Wang algorithm and the Metropolis and Gibbs sampler algorithms in the amount of time required for the data to become uncorrelated. Since we can't expect to get good statistics until after several

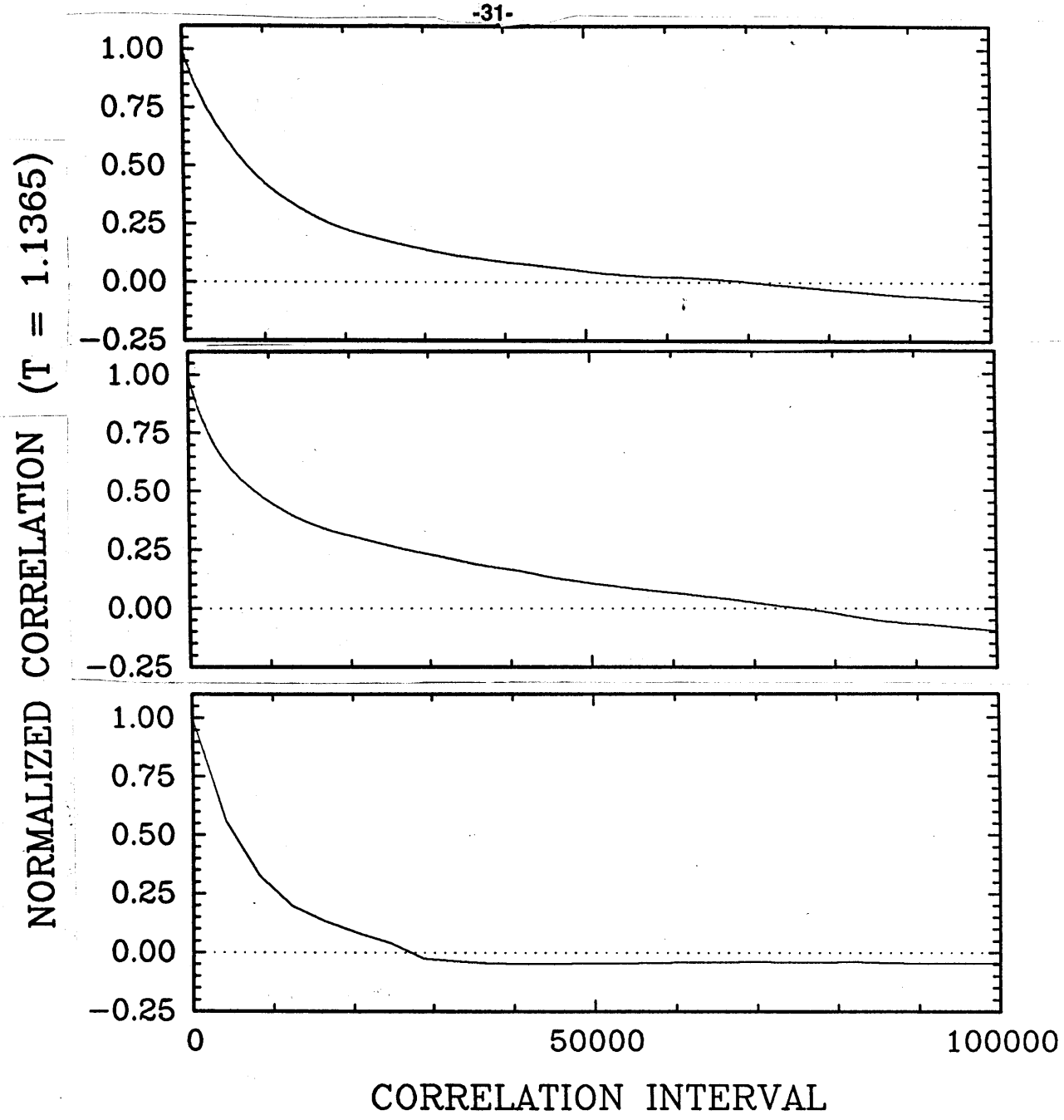

Figure 4-7: Temporal Correlation - Low Temperature

100 iterations per lattice site. Lattice size: 64x64. The graphs show the normalized temporal correlation of equilibrium energy over time. Algorithms from top to bottom: Metropolis, Gibbs sampler, Swendsen-Wang.

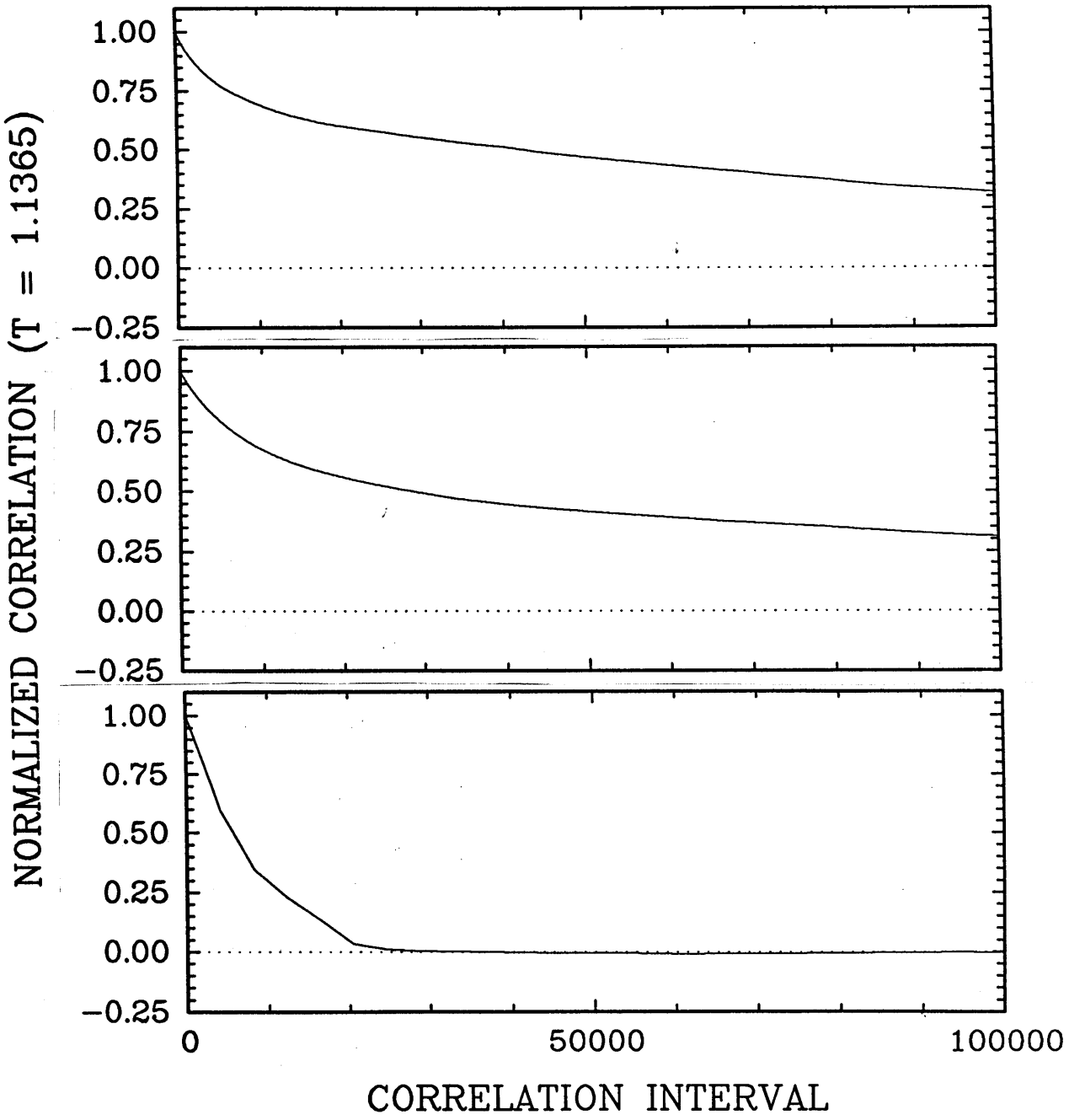

Figure 4-8: Temporal Correlation - Low Temperature

1000 iterations per lattice site. Lattice size: 64x64. The graphs show the normalized temporal correlation of equilibrium energy over time. Algorithms from top to bottom: Metropolis, Gibbs sampler, Swendsen-Wang.

times as many iterations as are required for the data to become uncorrelated, these figures effectively represent the speed with which the algorithms reach the desired results. The difference in algorithm efficiency is thus very pronounced at this low temperature, with the Swendsen-Wang algorithm arriving at a solution at more than twice the speed of the Metropolis and Gibbs sampler algorithms. The Metropolis and Gibbs sampler operate with near equal efficiency. These results reaffirm the comments made in the preceding section regarding low temperature performance of the three algorithms.

4.3.2 Critical Temperature

The next set of trials for this block of experiments were conducted near the critical temperature. Figure 4-9 on page 34 shows the results of the three algorithms for a run of one hundred iterations per site. Figure 4-10 on page 35 shows the same results for a run of one thousand iterations per site.

Once again the Swendsen-Wang algorithm shows substantial improvement in efficiency over the Metropolis and Gibbs sampler algorithms. This result is as expected since the Swendsen-Wang algorithm is still forming reasonably sized clusters and thus is still able to effect larger changes per iteration that the other two algorithms.

4.3.3 High Temperature

For the final set of trials for this block of experiments, the temperature was set well into the high temperature region. Figure 4-11 on page 36 shows the results of the trials for each of the three algorithms for a run of one hundred iterations per site. Figure 4-12 on page 37 shows the same results except for a run of one thousand iterations per site.

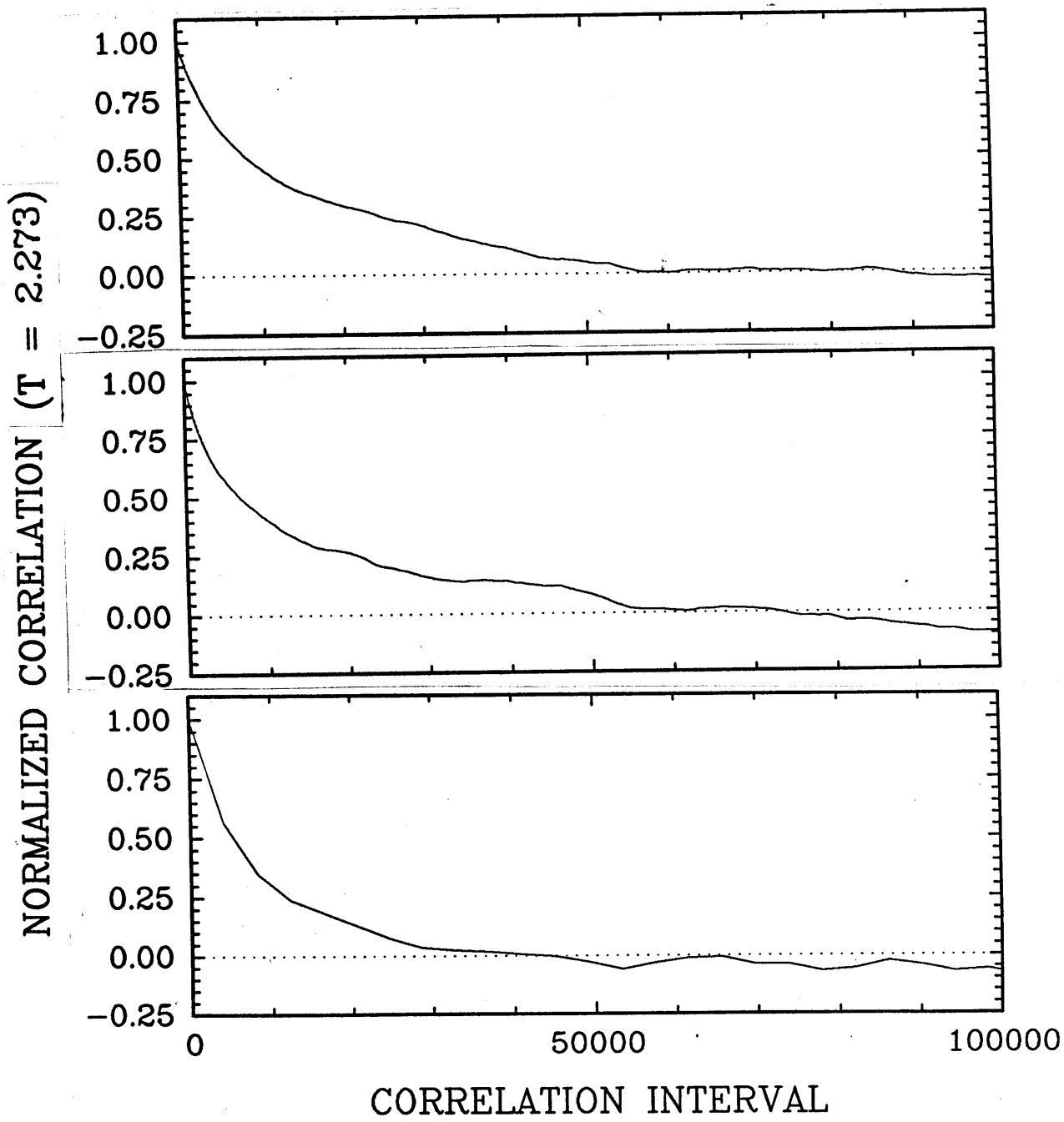

Figure 4-9: Temporal Correlation - Critical Temperature

100 iterations per lattice site. Lattice size: 64x64. The graphs show the normalized temporal correlation of equilibrium energy over time. Algorithms from top to bottom: Metropolis, Gibbs sampler, Swendsen-Wang.

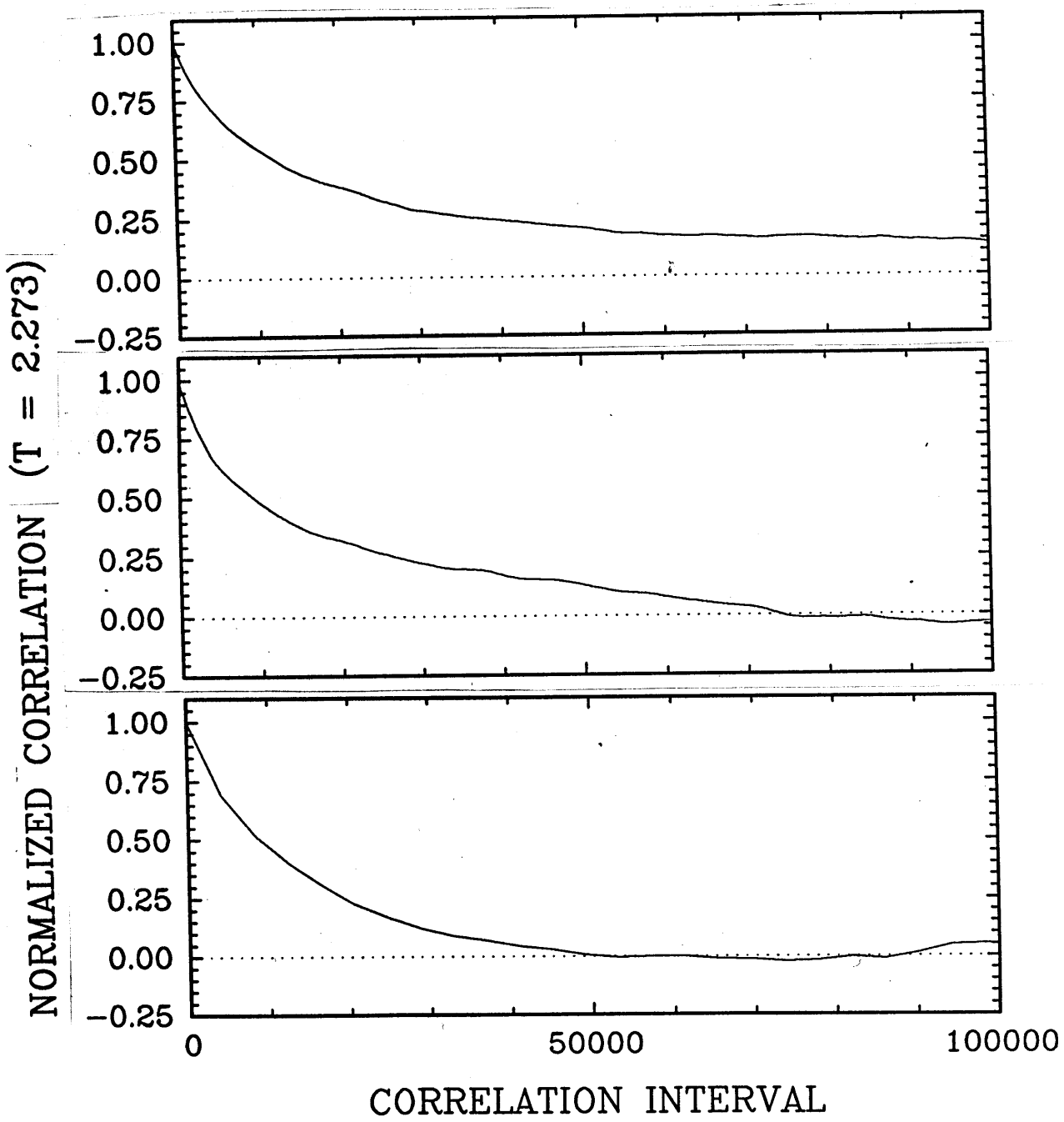

Figure 4-10: Temporal Correlation - Critical Temperature

1000 iterations per lattice site. Lattice size: 64x64. The graphs show the normalized temporal correlation of equilibrium energy over time. Algorithms from top to bottom: Metropolis, Gibbs sampler, Swendsen-Wang.

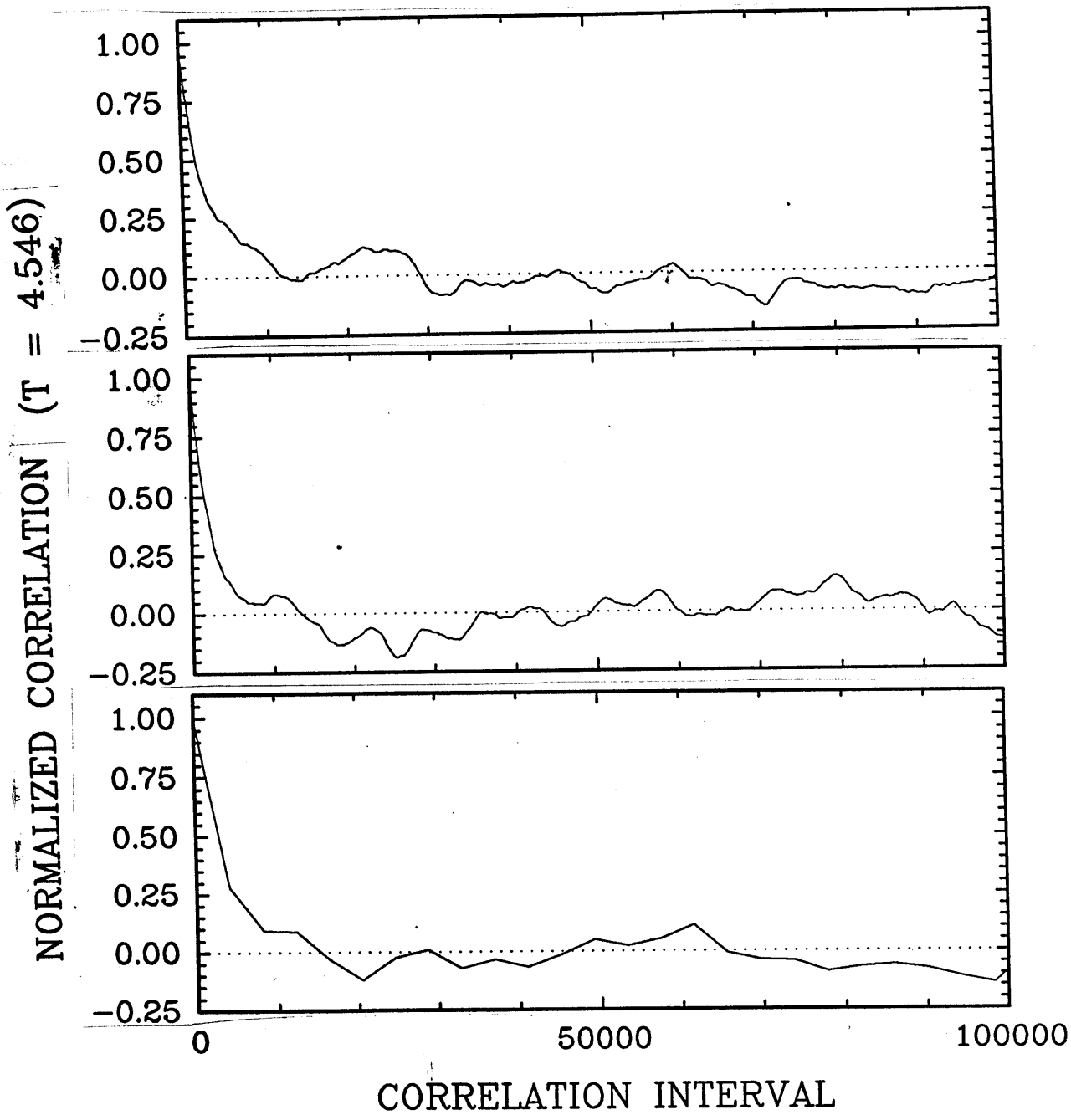

Figure 4-11: Temporal Correlation - High Temperature

100 iterations per lattice site. Lattice size: 64x64. The graphs show the normalized temporal correlation of equilibrium energy over time. Algorithms from top to bottom: Metropolis, Gibbs sampler, Swendsen-Wang.

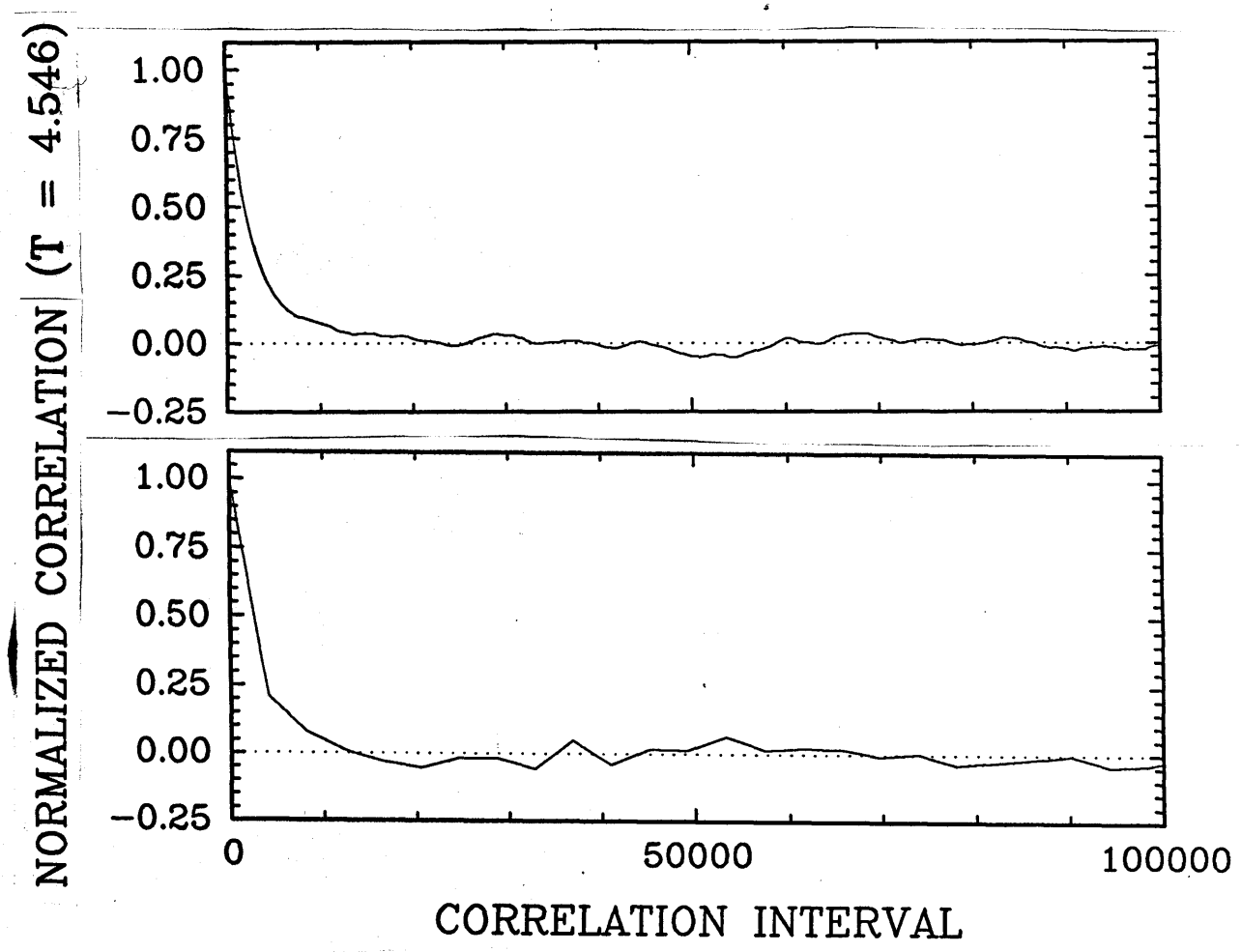

Figure 4-12: Temporal Correlation - High Temperature

1000 iterations per lattice site. Lattice size: 64x64. The graphs show the normalized temporal correlation of equilibrium energy over time. Algorithms from top to bottom: Metropolis, Gibbs sampler, Swendsen-Wang. Graphical data for the Gibbs sampler were not available.

The results obtained for this set of trials again confirm the observations made in the previous section. There is very little difference in efficiency between the three algorithms in the high temperature region. The data from all three of the algorithms becomes uncorrelated very quickly, and thus, all of the algorithms produce valid results in a very short time when operating in this temperature region.

4.4 Conclusion

Thus, the Swendsen-Wang algorithm outperforms the Metropolis and Gibbs sampler algorithms near and below the critical temperature. At higher temperatures the three algorithms are virtually equivalent in terms of efficiency and agreement with theoretical results. Since the Swendsen -Wang algorithm has nearly the same execution efficiency as the Metropolis and Gibbs sampler algorithms as discussed in the previous chapter (discounting the fact that the Gibbs sampler can be implemented using a parallel processing approach), the Swendsen-Wang algorithm has an overall efficiency equal to or greater than the other algorithms depending upon the operating temperature.

Chapter 5

Single Crystal X-ray Crystallography

5.1 Background

Determination of crystal structure through the use of X-ray crystallography is currently a long and complicated process. The crystallography data is essentially a fourier transform of the electron density of the crystal structure and only provides information about the amplitudes of the frequencies. Thus, the phase information is missing and must be reconstructed using some analytical method. The process at the present consists of several long and time consuming steps. What is desired is the ability to take the crystallography data and reconstruct the crystal structure in one step. The proposed method for accomplishing this goal is the use of a probabilistic model that incorporates the a posteriori information with an a priori distribution derived from a known solution to a similar crystal.

5.2 The Crystallography Model

The model selected for use was developed by Peter Doerschuk at the Laboratory for Information and Decision Systems at Massachusetts Institute of Technology and is similar in many respects to the Ising model discussed in the previous chapters. This approach discretizes the atomic locations and models the locations using a Markovian random field. Thus, the crystal is mapped on to a 3D lattice L with the points on the lattice defining the possible positions of atoms within the crystal. At each lattice point n, we define a

random variable ϕ_n which takes on the value 1 if the lattice site is occupied by an atom and 0 otherwise. The ϕ_n are modeled as a Markovian random field.

The a priori model chosen for this problem takes a fairly simple approach. The model simply requires that all atoms are covalently bonded to some other atom with a covalent bond of chemically appropriate length. The total number of covalent bonds to a given atom, the type of covalent bond, and the geometry of the covalent bonds are not specified. What we do specify is the appropriate ranges that the covalent bond lengths must fall within to be considered a proper bond length. Let n_{min} represent the minimum bond length and n_{max} represent the maximum bond length for a proper bond. Furthermore, define the neighborhoods of the Markovian random field as $N_n = \{m \; || m-n| <= n_{max}\} - \{n\}$. Therefore, $A_n = N_n - \{m \; || m-n| > n_{min}\}$ represents the set of all neighbors of site n that are too close to have a proper bond length with site n, and $B_n = N_n - A_n$ represents the set of all neighbors of site n that have a proper bond length with site n. Thus, the neighborhood for each site is separated into two concentric shells, one determining improper bond lengths (too close) and the other determining proper bond lengths.

Next, we consider the contribution of a site n to the total configuration energy. Since the physical interpretation is not clear as to what the contribution of an unoccupied site ($\phi_n=0$) is to the total configuration energy, we assume that there is no contribution to the configuration energy by unoccupied sites. Thus, we will consider only occupied sites or atoms $j = \{n \mid \phi_n=1\}$ on the lattice in the determination of total configuration energy. In

determining total configuration energy, we wish to penalize atoms that are improperly bonded (other atoms too close or no atoms at the proper bond lengths), and we wish to reward atoms that are properly bonded (only atoms at the proper bond lengths). The penalty and reward will take the form of a higher configuration energy contribution and a lower configuration energy contribution, respectively. Thus, we define the function

$$\xi_j = [1-\Pi_{m\varepsilon B_j}(1-\phi_m)][\Pi_{m\varepsilon A_j}(1-\phi_m)] \tag{5-1}$$

Thus, ξ_j indicates whether or not an atom is properly bonded. The β_0 and β_1 terms are parameters which weight the effects of the two bonding conditions upon the configuration energy. While these parameters are free to take on any value desired, they usually are of equal magnitude and opposite sign, with β_0 positive to effect a higher configuration energy and with β_1 negative to effect a lower configuration energy. For the contribution of atom *j* to the total configuration energy we have

$$u_j = \beta_0(1-\xi_j) + \beta_1\xi_1 \tag{5-2}$$

Hence, the total configuration energy is

$$U(\omega) = \sum_{<j>} u_j \tag{5-3}$$

where ω is the configuration on the lattice *L*. Finally, we define the invariant measure to be

$$\pi = \frac{1}{Z} exp(-U) \tag{5-4}$$

where Z is a normalizing constant.

Chapter 6

Swendsen-Wang Approach to the Crystallography Model

While Metropolis algorithms have been developed for the crystallography model described in the previous chapter, the development of an algorithm similar to the Swendsen-Wang algorithm for the Potts model would be a valuable asset in the crystallography project. This algorithm would borrow the notion of "bonds" from Swendsen-Wang and redefine it to fit the crystallography model. The overall idea would be to generate configurations in which bonded interactions maintain fixed low energy contributions. Thus, bonded interactions would have the same energy contribution in all configurations that they pass through.

6.1 Derivation of Algorithm

Following the reasoning of Swendsen and Wang (1987), we restrict the more general energy function given by equations 5-2 and 5-3 to

$$U = \beta \sum_{<i>} (1-\xi_i) \tag{6-1}$$

where the summation is over all occupied sites (atoms) and β is a parameter of the system. Hence, we penalize improperly bonded atoms by assigning it an energy of β while we neither penalize nor encourage properly bonded atoms. Using this quantity, we define the invariant measure to be

$$\pi = \frac{1}{Z} exp(-U) \tag{6-2}$$

Next, consider the energy contribution of an atom *I* and remove it from the configuration energy. Thus, we have

$$u_I = \beta \sum_{<i> \neq <I>} (1-\xi_i) \tag{6-3}$$

From this we can derive the restricted sums

$$\pi_I^{proper} = \frac{1}{Z} exp(-u_I)\,\xi_I \tag{6-4}$$

and

$$\pi_I^{imp} = \frac{1}{Z} exp(-u_I)(1-\xi_I) \tag{6-5}$$

such that

$$\pi = \pi_I^{proper} + exp(-\beta)\,\pi_I^{imp} \tag{6-6}$$

where π_I^{proper} represents the case where atom *I* is properly bonded, and π_I^{imp} represents the case where atom *I* is improperly bonded. Next, we introduce the term that corresponds to weighting the proper and improper bonding cases equally, thus generating a term independent of the state of atom *I*:

$$\pi_I^{ind} = \pi_I^{proper} + \pi_I^{imp} \tag{6-7}$$

Using equation 6-7, we can rewrite equation 6-6 so that

$$\pi = [1 - exp(-\beta)]\,\pi_I^{proper} + exp(-\beta)\,\pi_I^{ind} \tag{6-8}$$

The first term contains the restriction that atom *I* be properly bonded while the second term contains no restrictions whatsoever. Thus, the weighting factor for the first term can be interpreted as the probability, $p = 1 - \exp(-\beta)$, that the atom *I* be "bonded" in the same sense as in the Swendsen-Wang algorithm for the Potts model. Hence, if atom *I* is properly bonded in a given configuration in the crystallography model, then with probability *p* atom *I* will be properly bonded in the next configuration derived from the current configuration when traversing the Markovian chain.

6.2 Obstacles to Algorithm Development

Deriving a working algorithm for the crystallography model from the results obtained in equation 6-8 is difficult if not impossible. Two problems exist that hamper attempts to construct an algorithm. First, the nature of the energy function couples atoms together in a way that cannot easily be described mathematically. In other words, an atom's neighbors determine the energy contribution of that atom, but the individual effect of each neighbor upon the atom's energy contribution is not explicitly defined. Thus, it is not possible to uncouple restrictions demanding that an atom be properly bonded and restrictions demanding that other atoms must assume a state independent of any restrictions.

As an example, consider a certain crystallography algorithm in which the number of atoms is fixed. The methodology of this example algorithm is that every atom is visited without regard for order. If a visited atom is properly bonded, it is marked with a probability $p = 1 - \exp(-\beta)$; otherwise, it is left unmarked. Once all atoms have been visited, the atoms are removed from the lattice and set aside. Next, all unmarked atoms are placed on the lattice in entirely random sites. After all unmarked atoms have been placed, the marked atoms are placed one by one on the lattice in randomly determined sites with the restriction that the marked atom be placed in a site that will ensure that the marked atom will be properly bonded. At first glance, this algorithm appears to satisfy the invariant measure given in equation 6-8.

Now, consider a lattice that contains only two atoms that are positioned so that they are properly bonded. Indeed, since there are only two atoms, they could only be either both properly bonded or both improperly bonded. Upon performing one iteration of the example algorithm upon the lattice, say that one atom was marked and the other was left unmarked. So, the lattice is cleared and the unmarked atom is placed randomly upon the lattice. The marked atom is then randomly placed upon the lattice so that it is properly bonded. This means that it is placed with the proper bonding range of the unmarked atom. Thus, as long as at least one of the two atoms is marked, both will always end up being properly bonded. This result is contrary to the desired result since the independence of state required for unmarked atoms is compromised by the restriction that marked atoms be properly bonded. Essentially, it is impossible to enforce that certain atoms be properly bonded and also obtain independence of state for others.

The second problem is related to the first but focuses more on the crystallography model, the Potts model, and their relationship with the Swendsen-Wang algorithm. State within the Potts model and the interactions that determine configuration energy have a transitive nature. For example, if site *a* and site *b* have the same state, and site *b* and site *c* have the same state, then site *a* and site *c* have the same state. This is not necessarily the case with the crystallography model. Just because atoms *b* and *c* are both properly bonded to *a* does not mean that they are properly bonded to each other. The Swendsen-Wang algorithm works well with the Potts model since groups of sites with the same state always interact the same. Since states vary relative to spatial locations in the crystallography model, grouping of atoms into clusters is not feasible. Thus, due to the relative spatial nature of the crystallography model rather than absolute spatial nature of the Potts model, it is not possible to implement an algorithm analogous to the Swendsen-Wang algorithm in which large clusters of atoms are processed at the same time.

Chapter 7

Conclusion

Thus, while the Swendsen-Wang algorithm performs well for the Potts and Ising models, it is not possible to create an algorithm for the crystallography model that would be analogous to the Swendsen-Wang algorithm. However, this result does not entail considering future modifications that might allow an algorithm similar to the Swendsen-Wang to be developed for the crystallography problem. There are a number of potential changes that could be made to the crystallography model that look promising. First, the energy function could be modified to look at individual interactions between atoms rather than the aggregate effect that currently exists. Second, the entire lattice structure could be transformed onto another lattice structure on which the states look more like a Potts model. Hence, while the issue of adapting the Swendsen-Wang algorithm to the current crystallography model has been concluded unsuccessfully, the larger issue of adapting the Swendsen-Wang algorithm to the crystallography model is by no means closed.

References

Doerschuk P. Personal Communication (1985).

Geman S. and Geman D. "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images" IEEE Trans. Pattern Analysis and Machine Intelligence 6, 721-741 (1984).

Huang K. "Statistical Mechanics" John Wiley and Sons, Inc., 349-373 (1963).

Kindermann R. and Snell J. L. "Markov Random Fields and their Applications" Vol 1, Amer. Math. Soc. (1980).

Marroquin J. "Probabilistic Solution of Inverse Problems" Ph. D. Thesis, MIT Dept. of Elec. Eng. and Comp. Science (1985).

Metropolis N. et. al. "Equation of State Calculations by Fast Computing Machines" J. Phys. Chem. 21, 6 (1953) 1087-1092.

Swendsen R. and Wang J. "Nonuniversal Critical Dynamics in Monte Carlo Simulations" Phys. Rev. Letters 58, 2 (1987) 86-88.